A Data Fusion Algorithm for Mapping Sea-Ice Concentrations from Special Sensor Microwave/Imager Data

Kim C. Partington

Abstract—**Ice charts from the U.S. National Ice Center, Washington, DC, are compared to published algorithms for generating sea-ice concentrations from SSM/I data. The same ice charts, in a form that includes information derived only from RADARSAT, OLS, and AVHRR data, are used in an operational algorithm that effectively tunes a hybrid of the Bootstrap and NASA Team algorithms and principal components of the SSM/I data to the time and region associated with the ice chart. This "tuned" algorithm is then used to interpolate ice concentrations elsewhere in the ice chart where no cloud-free, high resolution, visible, infrared or active microwave satellite data are available. The algorithm is designed to operate in near real time to assist the ice analysts in their otherwise manually-intensive task of compiling ice charts for vessels operating in ice-infested waters.**

Index Terms—**Algorithm, data fusion, ice charts, passive microwave, sea-ice, SSM/I.**

I. Introduction

PASSIVE microwave data are unique in providing routine hemispheric-scale pictures of sea-ice conditions to the scientific and operational communities, and two decades of observations have now been accumulated by the Defense Meteorological Satellite Program (DMSP) special sensor microwave/imager (SSM/I) and its predecessor, the scanning multichannel microwave radiometer (SMMR). This is a sufficiently long time series to support studies related to climate-related events, such as the Antarctic Circumpolar Wave [1] and to extract information on systematic changes in ice conditions that may be related to climate change [2].

As well as being of scientific value, SSM/I data are also extremely important to the operational community, which has the task of supporting surface and subsurface ocean-going vessels with timely and regular information on ice conditions [3]. Ice analysts use SSM/I data to prepare ice charts for regular dissemination to the operational community, both in the U.S. at the National Ice Center (NIC), Washington, DC, and elsewhere around the world in areas where sea ice presents a hazard. SSM/I is important to the operational community because it provides wide and complete (cloud-free) coverage on a daily basis.

Although SSM/I data are important to the operational community, they do have limitations. SSM/I-derived sea ice products have coarse resolution (of the order of 25–50 km in the 19 to 37 GHz frequency range), so the data tend to be used after all higher resolution data sources have been exhausted. Other limitations are related to the algorithms used for converting the SSM/I brightness temperatures into ice concentrations. In contrast to active microwaves and visible and infrared satellite data, which are interpreted manually at the NIC, passive microwave data are converted to a geophysical product prior to use by the ice analysts. In general, these SSM/I algorithms do not make use of ancillary information so they are limited by the ambiguities inherent in passive microwave data. These ambiguities include confusion between weather artifacts, the presence of sea ice in areas of open water, and confusion between melting sea ice (for example indicated by surface melt-ponds) and areas of lower ice concentration [4]. The result of these algorithm limitations is that SSM/I data are less useful to the operational and scientific community than would be expected from the excellent coverage and the apparent transparency of clouds. In fact, with the exception of the location of the sea ice edge, SSM/I data are used by analysts at the NIC in a qualitative rather than quantitative manner in compiling ice charts.

The purpose of this paper is to show how operational ice charts provided by the NIC can be compared to sea-ice concentration algorithms and can form part of an operational data fusion algorithm. Operational ice charts provide the least ambiguous form of wide coverage reference information on ice conditions, as they are based on an expert manual synthesis of observations from a wide variety of sources. These include satellite observations, knowledge of sea ice and ocean climatology, considerations of continuity, weather predictions, model forecasts, ice charts from other ice centers, and, when available, airborne and ship observations.

II. U.S. National Ice Center Ice Charts

The NIC provides weekly ice charts for all ice-covered seas and achieves this by maintaining a group of qualified ice analysts, each with the responsibility for generating ice charts in one or more sectors of the Arctic and Antarctic. Some of the analysts are naval personnel who are assigned to the ice center for a period of 2–3 years, while others are civilian personnel who are not subject to limited assignments. The analysts are supported by a senior analyst who controls the quality of the final product. The ice analysts undergo formal training and follow a set procedure for generating ice charts, which ensures that high quality input data (high spatial resolution) takes precedence over lower

Manuscript received March 19, 1999; revised October 11, 1999.

The author is with the U.S. National Ice Center, Washington, DC 20395 USA (e-mail: kparting@hq.nasa.gov).

Publisher Item Identifier S 0196-2892(00)04135-8.

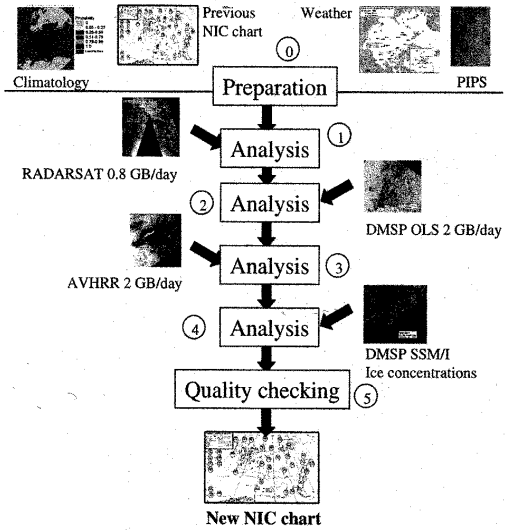

Fig. 1. Existing ice analysis procedure at the NIC, showing the use of SSM/I data processed using a conventional algorithm. Numbers indicate steps in the ice analysis procedure referred to in the text. The DMSP SSM/I ice concentrations are generated using both the NASA Team algorithm, which has been supplied by both the NOAA Environmental Prediction Center (NCEP) and the Fleet Numerical Modeling and Oceanography Center (FNMOC) and the CAL/VAL algorithm (FNMOC). To date, analysts have largely made use of the NASA Team product rather than the CAL/VAL product.

quality input data (low spatial resolution and/or cloud obscuration).

The ice analysis procedure at NIC is illustrated in Fig. 1. The analyst starts by compiling background information, including the ice chart from the previous week, digital climatology data [5], weather predictions, and output from the coupled ice-ocean forecast model, the Polar Ice Prediction System (PIPS) [6]. These data are used to provide the analyst with the best possible understanding of current, regional, background conditions from which to interpret the data (step 0, Fig. 1). The analyst then starts with the highest available resolution of data. On occasion, this may include some airborne or shipborne observations, but in general, these will not be available. In this case, the analyst starts with RADARSAT synthetic aperture radar (SAR) data. The analyst will map the region covered by RADARSAT data into regions of approximately "uniform" total and partial ice concentrations. In the event that there is ambiguity in the interpretation of RADARSAT data (for example, in terms of ice-water ambiguity in the signature), the analyst will use other data sources to resolve the interpretation, including weather fields, other satellite imagery, climatology, and the coupled ice-ocean model forecast. In practice, the ambiguity is almost always resolved with little residual uncertainty by the use of a combination of these other data sources. One or more of RADARSAT, AVHRR, and OLS data, along with knowledge of the ice climatology of the region, time of year, and weather predictions, is normally enough to resolve the presence of melt ponds or saturated snow on sea ice, for example.

Once there are no RADARSAT data available any longer, the analyst will use the next highest resolution data source, normally the operational line-scan system (OLS) data from the DMSP satellite program, which will be either visible or infrared data depending on the time of year and hemisphere. Having provided additional ice analysis based on OLS data, the analyst will then move to NOAA AVHRR data (step 3, Fig. 1). In the existing ice analysis procedure, two further steps remain. In the first step, DMSP SSM/I data are analyzed, having been translated into ice concentrations using the NASA Team algorithm, provided by the NOAA Center for Environmental Prediction (NCEP) (R. Grumbine, personal communication). These data are used to fill in the remaining areas of the new ice chart (step 4). As a final step, the ice chart is quality checked by a senior analyst (step 5, Fig. 1). An example of an ice chart, for the Sea of Okhotsk, is shown in Fig. 2 for the week of December 7–11, 1998.

In an ideal situation, the ice charts would be compiled on a daily basis and so each ice chart would be comparable, in its entirety, to a daily SSM/I product. In practice, NIC ice charts are compiled using data covering a period of up to 72 h and are provided on a weekly basis. However, the resulting ice charts, although representing a composite of up to three days of data, are not averages. Each closed region (polygon) in an ice chart represents information on ice conditions obtained from one data source and one date. It is planned that in the near future, it will be possible to relate ice chart polygons labeled with a particular ice concentration to both the source of data used to derive the ice concentration in that particular region, as well as the actual date [3]. This will be achieved through the use of links in the ice chart database.

There are, of course, errors in the ice concentrations obtained manually by the analyst. To begin with, the estimate of ice concentration is subjective and is specified as a range of values to reflect this. In general, a range of ice concentrations is specified (for example, if the analyst estimates 8/10 ice concentration, they will typically specify a range of 7/10 to 9/10). However, toward the upper and lower limits of ice concentration, the errors are typically smaller and so the range of values can be smaller (pack ice is often estimated as 9 to 10/10 ice concentration and land-fast ice is estimated with some confidence as 10/10 ice concentration). In general, however, analysts tend to estimate an error in the estimate of ice concentration of about $\pm 1/10$. The errors will also include those related to resolution of the data being used. If ice is not resolved within a pixel of a certain type of image data, then no ice will be detected. Thus, it is possible that very low ice concentrations where the floe sizes are small, will be underestimated in NIC ice charts if the ice is highly dispersed. Furthermore, the ice chart does represent a composite of information on ice conditions covering a period of about 72 h and so there will be errors made in assuming that the data are associated with a particular day.

Despite these limitations, the strength of the ice chart analysis procedure is that it is a form of manual data assimilation that takes into account:

1) physically realistic situations (e.g. continuity);
2) trends in ice conditions from the previous analysis;
3) a wide variety of data (satellite microwave, IR/VIS);
4) forecast ice conditions;

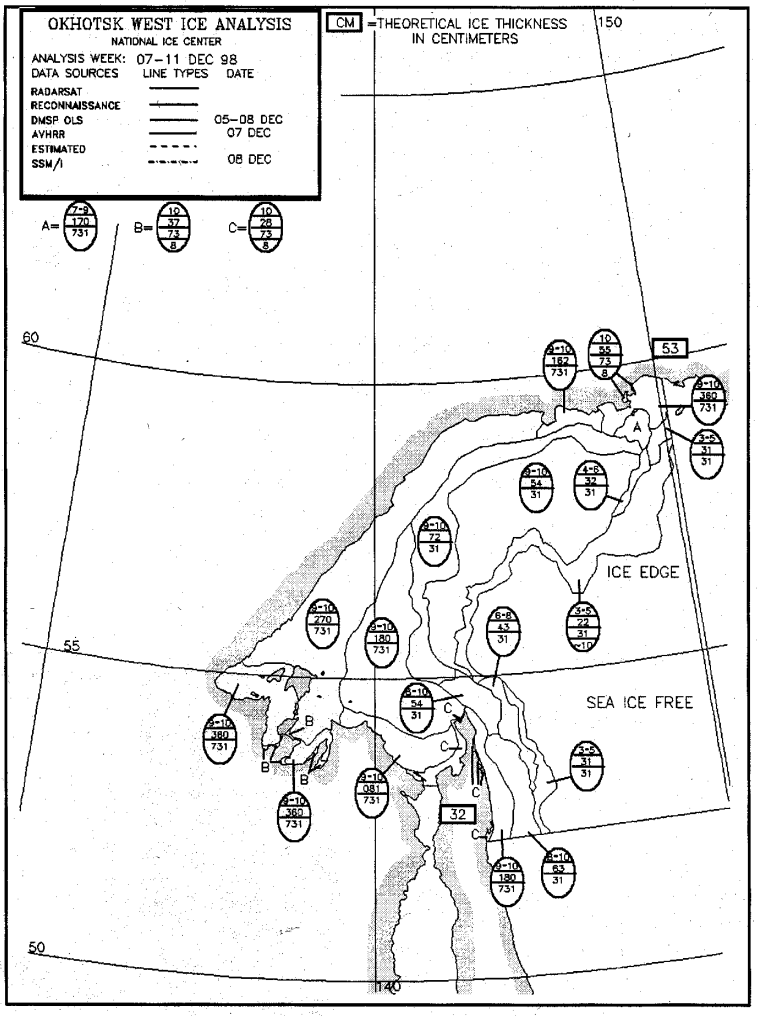

Fig. 2. Example of an NIC ice chart, disseminated to users on a weekly basis for all ice-covered seas. This example shows the western Sea of Okhotsk for December 7–11, 1998. The symbols are called "egg codes," and they are explained at the U.S. NIC web site (http://www.natice.noaa.gov). They provide information on total and partial ice concentrations. A partial ice concentration is the concentration of a particular ice type.

5) weather predictions;
6) digital climatology (statistically likely conditions);
7) knowledge of regional geography (ocean currents, etc.);
8) any reconnaissance information.

No statistics are available to provide a quantitative assessment of the accuracy of the ice charts. This would require wide coverage and high accuracy information on ice concentrations derived using data that are not used within the ice analysis procedure. Thus, we cannot claim that the ice charts represent any sort of absolute reference against which to test algorithms. Some limited coverage evaluations of the ice charts could be carried out, for example, using a SAR data set aside from the ice analysis procedure. However, even then, the evaluation would not be extendable in terms of conclusions to other seasons and regions. Thus, in describing the new SSM/I algorithm in this paper, no independent evaluation is provided, and this represents a limitation of the technique.

III. Limitations of Conventional SSM/I Ice Concentration Algorithms

It was shown in Fig. 1 that SSM/I is the last data source used to compile ice charts within the ice analysis procedure, primarily because it has the lowest spatial resolution of satellite data. However, SSM/I is also appropriately analyzed last, because it has some significant biases, particularly in the marginal ice zone and during summer when accuracies are much reduced (compared, for example, to freezing conditions in the central Arctic). On the other hand, SSM/I data are always available and so it remains an invaluable source of data, especially in the southern hemisphere where satellite SAR data are currently unavailable on a routine, operational basis.

SSM/I algorithms have been based largely on simplified versions of radiative transfer models of the surface and atmosphere, together with a mixing model that assumes that the footprint of

TABLE I
A SELECTION OF EVALUATIONS OF NASA TEAM, BOOTSTRAP AND CAL/VAL ALGORITHMS. * INDICATES SIMPLIFIED VERSION OF CAL/VAL CALLED THE AES/YORK ALGORITHM [10]. "MEAN" INDICATES THE MEAN DIFFERENCE, WITH A POSITIVE VALUE INDICATING THAT THE SSM/I ALGORITHM INDICATED MORE ICE THAN THE EVALUATION DATA. STD INDICATES STANDARD DEVIATION AND RMS INDICATES ROOT MEAN SQUARE (RMS) DIFFERENCE

Reference	Conditions, Hemisphere	Evaluation method	NASA Team accuracy	Bootstrap accuracy	CAL/VAL accuracy	Non-specific algorithm evaluation
Steffen et al. [4]	Freezing North and South	Review article	N/a	N/a	N/a	In general 3%-10% error
Comiso et al. [13]	Freezing, North	Landsat	Mean -8.2% Std 8.8%	Mean -6.1% Std 9.0%	N/a	N/a
Smith [11]	Freezing, North	AVHRR	Rms 7.2%	Rms 8.3%	Rms 6.9%*	N/a
Emery et al. [14]	Freezing, North	AVHRR	Mean -5.7% Std 8%	Mean -5.3% Std 9.7%	N/a	N/a
Steffen and Schweiger [15]	Freezing, North	Landsat	Mean –3.6% Std 6.6%	N/a	N/a	N/a
Cavalieri et al. [12]	Freezing, North	Airborne sensors	Mean -2.4% Std 2.4%	N/a	N/a	N/a
Emery et al. [14]	Melting, North	AVHRR	Mean -3.1% Std 8.8%	Mean 4.8% Std 9.8%	N/a	N/a
Steffen et al. [4]	Melting, North and South	Review article	N/a	N/a	N/a	In general 10%-20% error

the sensor contains only open water and sea ice. These models include

1) the NASA Team algorithm and modification for thin ice ([7], [8]);
2) the Bootstrap algorithm [9];
3) the CAL/VAL (modified version of AES-York) algorithm [10].

Other algorithms have also been developed (see [11]). The differences lie in the choice of frequency and polarization combinations and in the precise technique used to estimate concentration from the distribution of points in the relevant feature space.

The techniques listed here all made use of one or both of the 19 and 37-GHz channels on SSM/I (with the 22 GHz channel sometimes used for data quality checking), but they generate significant differences in predictions of ice concentration. Steffen *et al.* [4] evaluated seven algorithms and concluded that total ice concentration can be estimated, at best, with an accuracy of 5–10% during nonmelt conditions and 10–20% at other times. It has been estimated that, during the winter months, the NASA Team algorithm has an accuracy of -3.6% $\pm 6.6\%$ [12]. More recently, Comiso *et al.* [13] found that both the Bootstrap and NASA Team algorithms persistently underestimated ice concentrations when compared to AVHRR, LANDSAT, and SAR data. Table I summarizes some of the differences found by various authors.

To augment this evaluation, a comparison of just two algorithms is provided in Fig. 3 for a random day in the Arctic (December 9, 1998). Table II shows statistics that summarize differences between these two algorithms in various regions across the Arctic and Antarctic.

It can be seen that differences in estimated ice concentration are substantial and vary significantly from region to region, with the greatest differences in the Sea of Okhotsk and the smallest differences in the central Arctic. These differences are larger than would be expected given the figures in Table II and indicate the extent of the uncertainty involved in using SSM/I as a tool to aid operational ice analysis, particularly in the marginal ice zone. Indeed, the figures in Table II suggest that previous estimates of errors tend toward the optimistic, unless the estimates are restricted to pack-ice in freezing conditions.

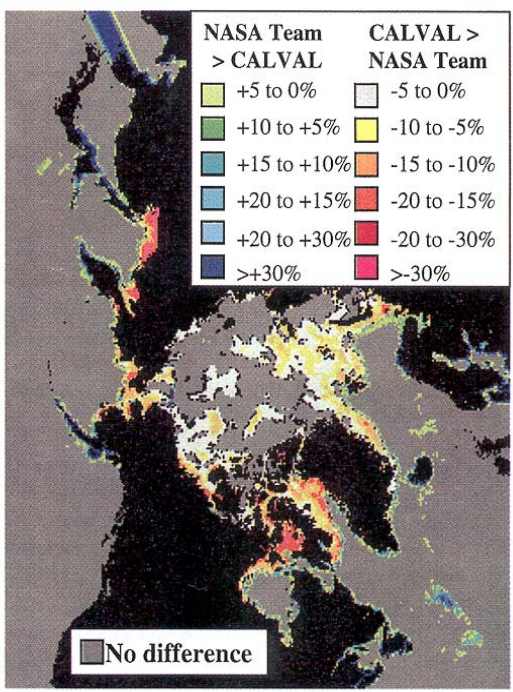

Fig. 3. The difference between the CAL/VAL and NASA Team algorithms for December 9, 1998, in the northern hemisphere, in terms of percent ice cover.

TABLE II
STATISTICS RELATED TO DIFFERENCES IN ICE CONCENTRATIONS ESTIMATED USING THE NASA TEAM AND CAL/VAL ALGORITHMS FOR DECEMBER 9, 1998. A POSITIVE SIGN INDICATES THAT THE CAL/VAL ALGORITHM SHOWS MORE ICE THAN THE NASA TEAM ALGORITHM (IN THE RIGHTMOST COLUMN, THE NEGATIVE SIGN INDICATES THAT THE NASA TEAM ICE EDGE EXTENDS FURTHER EQUATOR-WARD THAN THE CAL/VAL ICE EDGE). THE NUMBER OF PIXELS USED FOR EACH REGION IS BROADLY RELATED TO THE SIZE OF THE REGION, WITH COASTAL PIXELS NOT USED

Region (9 December 1998)	Season	Mean concentration difference (%)	R.m.s. concentration difference (%)	Mean 0% ice concentration edge position difference (km)
Central Arctic	Winter	2.6	3.34	-71.0
East Greenland	Winter	8.8	5.4	-23.0
Baffin Bay	Winter	13.5	6.2	-23.0
Sea of Okhotsk	Winter	29.0	8.9	-44.0
Ross – Bellinghausen Seas	Summer	14.5	5.3	-24.0
Weddell Sea	Summer	15.3	6.2	-21.0

To understand why these differences can be so large and vary systematically from region to region, an analysis of the different sources of error is required. A brief review of the errors is provided here. For more detail, the reader is referred to referenced publications.

A. *Tie-Point Variations*

Tie-point related errors result from natural variations in brightness temperatures associated with inhomogeneities within individual surface types. SSM/I algorithms make use of reference brightness temperatures associated with two ice types (first-year ice and multiyear ice in the northern hemisphere) and open water. These values are known as tie points. These tie points are used to establish definitions for 100% concentrations of the two ice types and open water. However, all natural surfaces have variability in their emissivity characteristics and physical temperatures. In the NASA Team algorithm, for example, the open water tie-point is biased toward calm water conditions, and so areas of rough water can be given an erroneously nonzero ice concentration [12]. Furthermore, Steffen and Schweiger [15] found that, in the northern hemisphere, the NASA Team algorithm global tie point for first-year ice tends to be higher than locally determined first-year ice tie points, which provides one explanation as to why ice concentrations in heavy pack are generally underestimated with the NASA team algorithm. Comiso *et al.* [13] found that ice concentrations derived from the NASA Team and Bootstrap algorithms generally varied about by between 1% and 2%, purely in response to natural differences in the reference emissivities of ice and open water.

Natural variations in tie points can be related to natural variations in a number of characteristics of sea ice and open water. The open water brightness temperature, for example, will vary according to physical temperature, wind speed, foam, and atmospheric effects. Sea ice brightness temperatures will vary according to physical temperature, atmospheric effects, snow cover characteristics, sea ice salinity, density, and roughness. To take one example, snow cover has been shown to have an effect on passive microwave signatures at 37 GHz and above (Fig. 4, also [9], [13], [16], [17]). The effect is not only one of influencing the signature through variations in snow thickness, but also influencing the signature through differences in granularity,

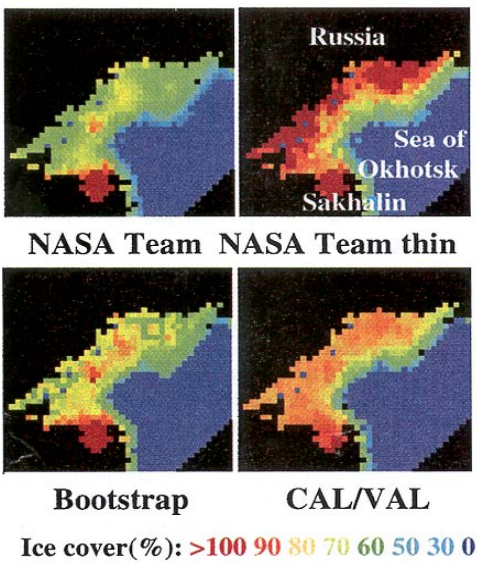

Fig. 4. Algorithm predictions of total ice concentration for the western Sea of Okhotsk on December 9, 1998. The corresponding ice chart, compiled using nearly cloud-free OLS data, is shown in Fig. 2. Total concentrations indicated by the ice chart are generally above 9/10 within the ice edge and partial concentrations of new and young ice types make up a significant percentage of the total concentration.

brine content (if any), and the presence of layering. Ice layers in snow influence horizontal polarization rather than vertical polarization and so have a particularly significant effect on algorithms that employ the horizontal polarization at 37 GHz. The Bootstrap and CAL/VAL algorithms make use of the 37 GHz horizontal polarization and so are likely, on a priori grounds, to be more sensitive to variations in snow conditions than the zz algorithm. There do not appear to have been any studies that attempt to quantify the effect of snow thickness on ice concentration estimates, but it seems likely that variations in snow conditions will be equivalent to a local change in the tie-point that can translate into a modest error in ice concentration. There are a range of other geophysical characteristics that influence the

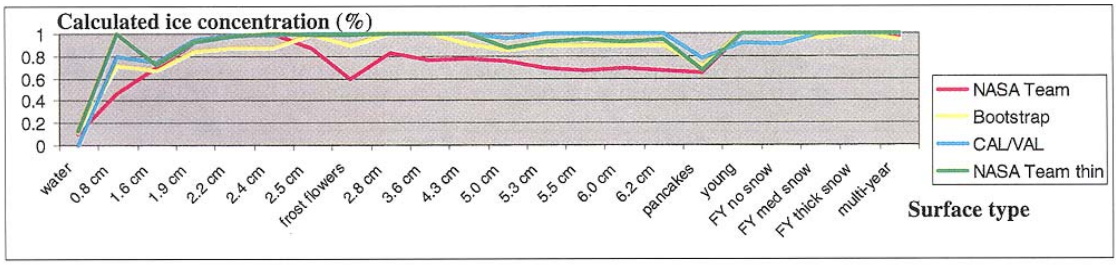

Fig. 5. Algorithm predictions of total ice concentration as a function of ice types (fraction ice cover), with brightness temperatures derived from controlled observations. Observations of new ice (with thicknesses indicated in cm) are derived from Wensnehan *et al.* [27], recorded over growing saline ice. Observations from frost flowers, pancakes, young, and multiyear ice were recorded by a shipborne radiometer sensor operating during the North Water experiment, Baffin Bay, May 1998 (Asmus, personal communication, 1999). First-year ice data with different snow thicknesses were taken from Grenfell *et al.* [28]. Given that the footprints in each case contained 100% of the labeled surface type, the "true" total concentrations are 0% for open water and 100% for all ice types. All observations correspond to freezing conditions.

tie-points and can be discussed, but instead the reader is referred to [4], [12] and [13] for more details.

Steffen and Schweiger [15] and others have found that the use of local tie points can (in some circumstances) halve the RMS difference in ice concentrations derived from SSM/I and visible/infrared sensors (AVHRR and Landsat), and the use of local tie points is particularly important in the Arctic. However, there are practical difficulties in implementing an SSM/I ice concentration algorithm that is able to make use of local tie points in any sort of routine manner.

B. Surface Temperature-Related Errors

Some researchers have noted that sea ice concentration can be correlated to surface temperature, and this is a form of error known as geophysical cross talk. The use of radiance ratios in the NASA team algorithm enables variations in surface temperature to be removed from having an effect on ice concentration to first order (as the brightness temperature, for a particular frequency and polarization, is equal to the temperature of the ice multiplied by the emissivity). However, there may be a small ice temperature effect on the (spectral) gradient ratio used in the NASA Team algorithm, as the temperatures sensed by the two different frequencies may be different due to different radiation penetration depths. Other algorithms are probably more sensitive to this error than the NASA Team algorithm. The Bootstrap algorithm shows some sensitivity in the Southern Ocean where there is a 2 to 5% error in ice concentrations related to ice temperature variations (estimated as 0.9% per Kelvin snow–ice interface temperature change) [13]. A result is that the Bootstrap algorithm can underestimate ice concentration when temperatures are on the low side (for example, -20° C). The CAL/VAL algorithm appears insensitive to ice temperature changes, primarily because it is designed for the ice margin, where surface temperatures are close to freezing point, and the algorithm tends to saturate (provide ice concentrations in excess of 100%) in the interior of the ice pack.

C. Surface Melt-Related Errors

These errors cause most SSM/I algorithms to underestimate ice concentration. Errors in ice concentration tend to double to 10–20% for algorithms during summer, and the errors are systematic [4]. In a sense, this is not an "error," as it reflects the fact that SSM/I classifies surface water as "open water." Russian drift stations show that, in mid July, as much as 45% of the ice may be covered by melt ponds [18], and this will be reflected in lower estimates of ice concentration from passive microwave data. As the sensor cannot penetrate beneath melt ponds, this problem cannot be solved in any simple manner. The problem will affect measurements as long as there is liquid water in significant quantities, such as in overlying snow cover.

D. Thin Ice-Related Errors

Errors related to the presence of new and young ice types are not explicitly accounted for in the existing operational SSM/I ice concentration algorithms. Thin ice types have passive microwave signatures that are distinct from both open water and first-year sea ice and are not accounted for explicitly in conventional SSM/I algorithms. Thus, in general, the application of a conventional SSM/I algorithm to an area of 100% thin ice concentration will tend to estimate it as some percentage of open water cover. In freeze-up conditions, there will always be a tendency for ice concentration to be underestimated as a result of this effect. Fig. 5 shows sea ice concentrations predicted for the Sea of Okhotsk on December 8, 1998, where a high concentration of new and young ice types (indicated by the ice chart in Fig. 2) creates major discrepancies in the predictions of ice concentrations from the different algorithms. A tendency of the NASA Team algorithm to significantly underestimate ice concentration in the Sea of Okhotsk has been noted by [19]. The differing sensitivities of the algorithms to thin ice types is confirmed in Fig. 4, which shows apparent ice concentration as estimated using the four algorithms listed above with brightness temperatures obtained from controlled observations from known ice types. It can be seen that the algorithms have varied responses to the presence of thin ice, pancakes, and frost flowers.

The explanation for this problem is that the emissivity of new ice can range from that of open water to first-year ice (0.45 to 0.92). Steffen and Schweiger [15] found that the presence of nilas and young ice could create an underestimate of ice concentrations from SSM/I of 9% (with global tie points) and 4% (with local tie points), using the NASA Team algorithm. The results shown here indicate that the errors can be much larger in areas where thin ice types predominate. A thin ice formulation

of the NASA team algorithm helps to correct for these ice types ([8] and Fig. 5), but the algorithm cannot deal simultaneously with multiyear ice, first-year ice, and thin ice.

E. Atmospheric Transmission Errors

These errors are caused by the susceptibility of passive microwave sensors to cloud liquid water and water vapor, integrated along the path length of the radiation. These errors are currently impossible to predict reliably, so SSM/I algorithms make use of weather filters to reject ice concentrations over open ocean. Ice concentrations can be affected by both integrated atmospheric liquid water and water vapor content. Oelke [20] has used radiative transfer modeling to assess the impact of weather systems on ice concentrations derived using the NASA Team algorithm. He found that cloud liquid water content and water vapor both artificially increase total ice concentrations, but decrease multiyear ice concentration. Near surface winds tend to increase ocean emissivity and thus add to spurious ice concentrations over the open ocean. Sensitivity to the atmosphere has been the main reason why the 85-GHz channel has not been as popular for retrieving ice information as its high spatial resolution would suggest ([21], [22]). Researchers are currently working on methods to account explicitly for the atmosphere in SSM/I algorithms [23].

F. Resolution-Related Errors

These errors represent uncertainty in the location of data. The CAL/VAL algorithm, for example, uses the 37-GHz channel alone to estimate ice concentrations at the ice edge, and this results in an ice edge that is generally at least 20 km poleward of the ice edge generated by the NASA Team algorithm (which uses the 19-GHz channel). This is a season-and-hemisphere independent result that reflects the different resolutions associated with the 19 and 37-GHz channels, which have an effective field of view (along-track, across-track) that is (69, 43 km), at 19 GHz and (37, 28–29 km) at 37 GHz (Table II). The finite resolution also explains why the rule-of-thumb for the NASA Team algorithm is that the actual ice edge corresponds to the 15% ice concentration contour [12].

In summary, there are significant deficiencies in the conventional SSM/I algorithms that make the search for climate signals and operational use problematic. The results presented here suggest that the errors are larger for some algorithms than has been suggested by other studies (at least on a regional basis). A mean difference of 29% in ice concentration for the Sea of Okhotsk from two algorithms, with maximum differences in excess of 50%, is extremely large. In order to improve the estimation of ice concentrations from SSM/I for operational ice monitoring, a different approach is needed than that provided by conventional SSM/I algorithms.

IV. Data Fusion Algorithm for Deriving Ice Concentrations from SSM/I

The issues discussed above limit the utility of conventional algorithms for operational ice monitoring and for scientific applications and it has become increasingly clear that substantial improvements in the quality of information derived from SSM/I can only be achieved by bringing in ancillary information. The conclusion from Steffen et al. [4] was that data assimilation/fusion and artificial intelligence methods (possibly making use of decision making to operate a hybrid algorithm) offer the greatest promise for resolving ambiguities in the passive microwave algorithms.

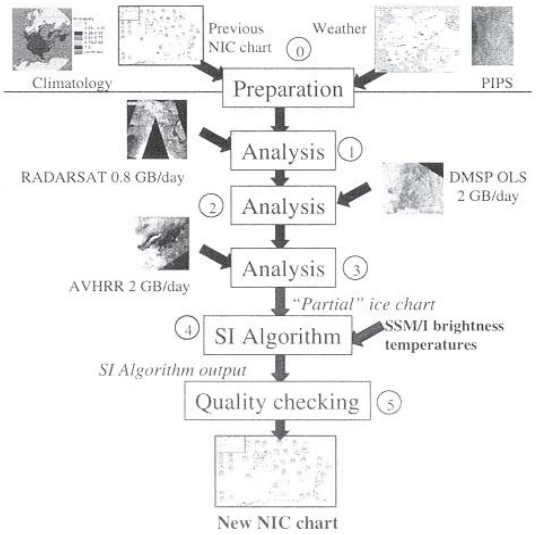

Fig. 6. Proposed new ice analysis procedure, showing the formation of a partial ice chart and its use in the SSM/I interpolation (SI) algorithm.

Some researchers have already begun to explore the possibility of retrieving ice information from SSM/I in combination with other data or models, as a way to reduce the ambiguities of weather effects, etc. These include the combination of AVHRR and SSM/I to classify sea ice, open water, land, and cloud [24], the constraint of SSM/I ice concentrations using SAR data [25], and the use of a physical model to constrain ice concentration estimates from SSM/I over time [26]. A data assimilation approach to the problem has particular potential, and in the long term appears to be the most promising way forward. In the shorter term, using other forms of data to constrain the SSM/I algorithm in a form of data fusion is easier to implement. However, if the data that are used to constrain the SSM/I are themselves ambiguous, as in the case of AVHRR (through cloud-ice ambiguities) and SAR (through water-ice ambiguities), it can be difficult to design an algorithm that is operationally practical.

The algorithm proposed here, provisionally called the SSM/I interpolation (SI) algorithm, attempts to compensate for limitations in the conventional algorithms and biases inherent in the sensor, by tuning the algorithm to ice concentration information available in NIC ice charts. The aim is that this would be a near-real time procedure in which the ice analyst routinely makes use of a tuned SSM/I ice concentration algorithm during the compilation of ice charts. The proposed context of the new algorithm is illustrated in Fig. 6, which should be compared with Fig. 1. The SI algorithm would replace the stage in the analysis procedure, where a conventional SSM/I algorithm is employed. The analyst would generate an ice chart with closed regions at step 3 in the analysis procedure, and all areas that had not been

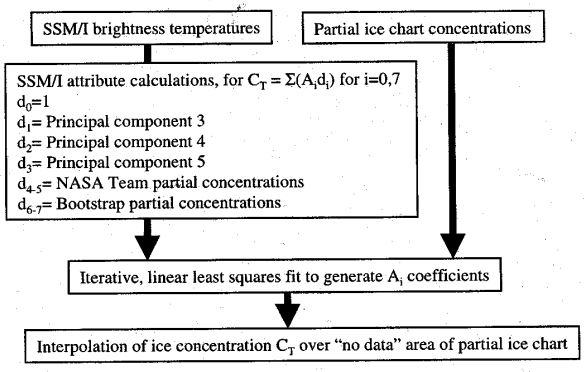

Fig. 7. Flow diagram showing details of the SI algorithm. The number associated with the principal component (3, 4, or 5) refers to the level of overall variance explained, so that 3 would indicate the component with the third highest level of variance. The principal components are calculated using data from the region covered by the ice chart and do not include land.

assigned ice concentrations at that point would be given a "no data" label. This is the so-called "partial" ice chart, which is based solely on reconnaissance, RADARSAT, and cloud-free, AVHRR and OLS data. The analyst would then run the SI algorithm on the partial ice chart, and it would optimally match calculated SSM/I attributes to the ice concentrations in the partial ice chart using the procedure illustrated in Fig. 6. The calculated model coefficients would then be used to "interpolate" ice concentrations for the "no data" regions remaining in the partial ice chart. The analyst would then use these interpolated ice concentrations from SSM/I to complete the ice chart.

This approach has the following advantages.

1) The algorithm is tuned as frequently as ice charts are generated (weekly in most regions and twice weekly in some).
2) The algorithm is tuned to the region being analyzed, effectively taking into account regional anomolies in salinity, snow conditions, etc. This is equivalent to generating location-specific and time-specific tie points.
3) The tuning parameters can be used to help define a regional and temporal database for future and retrospective processing of passive microwave data.

The SI algorithm is a simple, statistical least squares procedure that takes attributes derived from the SSM/I brightness temperatures and generates linear mapping coefficients that allow the SSM/I attributes to be mapped to total ice concentration, with the latter extracted from the so-called partial ice charts. The equation is defined in Fig. 7. A critical element of the algorithm is the choice of attributes, as inappropriate attributes will make the mapping function inefficient. The attributes selected for implementation include the partial ice types calculated by the NASA Team and Bootstrap algorithms and selected principal components of the SSM/I data. The rationale for including these attributes is as follows.

1) The NASA Team and Bootstrap algorithms provide independent estimates of two partial ice type concentrations ([7], [9]). In the northern hemisphere these correspond predominantly to first-year and multiyear ice. The

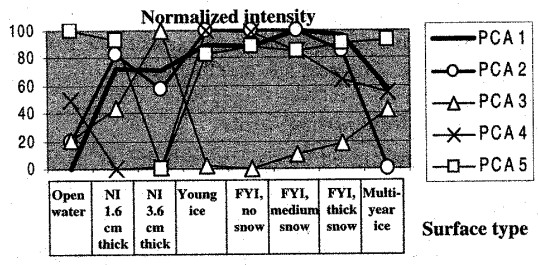

Fig. 8. Principal components of the SSM/I data applied to controlled sea ice and open water observations (a subset of those in Fig. 5), where the eigenvectors are generated using principal component analysis from the entire Arctic (excepting land masses) for December 9, 1998. The component intensities are normalized from 0 to 100.

NASA Team algorithm explicitly provides equations for these two partial ice type concentrations. The Bootstrap algorithm does not provide this information explicitly, but Comiso [9] shows that the end clusters along the line corresponding to 100% ice concentration in both polarization and frequency space (clusters A and D in Fig. 9 in [9], for example) represent first-year and multiyear ice in the northern hemisphere. Hence, it is possible to calculate the proportions of these two ice types in the Bootstrap algorithm by assuming that these two end points represent 100% concentrations of each ice type. The partial ice types from both algorithms, added together as components of a linear function, are expected to be efficient measures of the concentration of ice of young age and older (Fig. 4). However, note that within the SI algorithm, no assumption is made about what particular ice types these partial concentrations refer to. The fact that these partial concentrations are multiplied by linear coefficients should enable the SI algorithm to compensate, to some degree, for underestimation of ice concentration in the summer melt period. The NASA Team and Bootstrap partial concentrations were selected because other algorithms are less well evaluated or do not provide as much information. The CAL/VAL algorithm, for example, generates ice concentration predictions that saturate at 100% over most of the ice extent, so it was likely that it would be inefficient as an attribute for the SI algorithm. However, additional or different algorithm concentrations could be added to the SI algorithm if required.

2) The lower variance principal components of the SSM/I data tend to be related to the presence of thin ice. This has been demonstrated by Wensnehan *et al.* [27] and Grenfell *et al.* [28], following earlier principal components analysis [29]. Hence, these attributes are included to counter the known tendency of the NASA Team and Bootstrap algorithms to underestimate ice concentrations in areas of thin ice. Fig. 8 shows third, fourth, and fifth principal components of data from December 9, 1998 are all strongly related to the presence of new ice. The third, fourth, and fifth principal components of the data (in order of decreasing variance) are included as a way of picking out conditions that the NASA Team and Bootstrap algorithms fail to detect, including the presence of thin ice

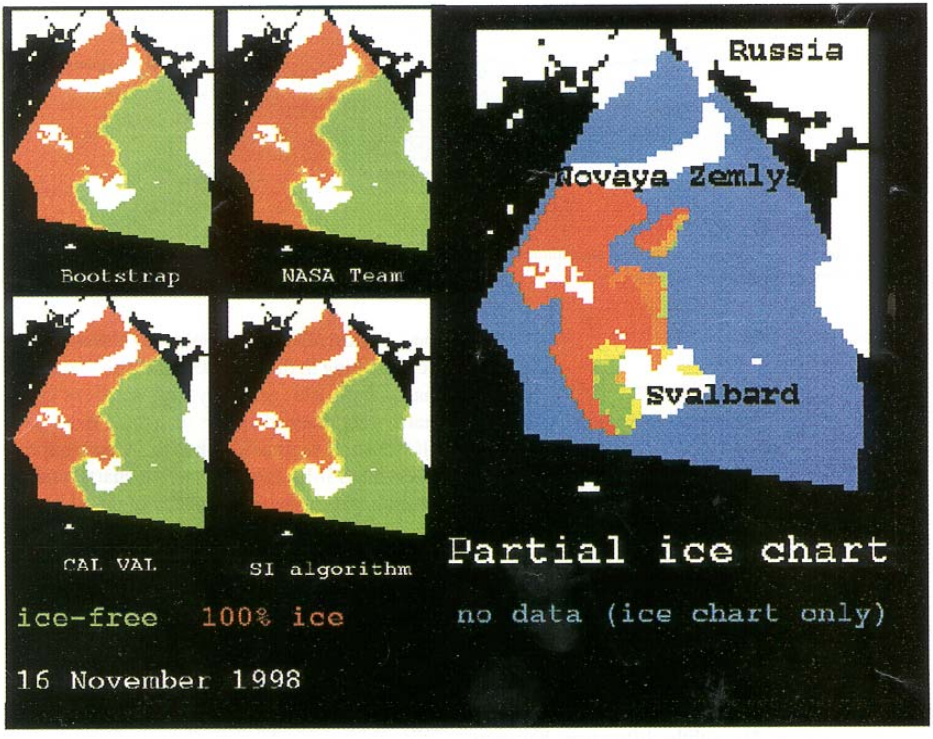

Fig. 9. Maps showing predictions of ice concentration from the NASA Team, Bootstrap, CAL/VAL, and SI algorithm for the Barents Sea. Also shown (right), is the partial ice chart used to "train" the SI algorithm. Blue indicates no data, red indicates 100% ice concentration, and green indicates 0% ice concentration. Land is shown in white. North is approximately toward the lower left corner of each map.

TABLE III
CONVERSIONS FROM RANGES OF ICE CONCENTRATIONS PROVIDED IN ICE CHARTS TO DISCRETE ICE CONCENTRATION VALUES

Ice concentration range (tenths)	Assigned value	Estimated error
0 to 1	0.05	±0.05
0 to 2	0.1	±0.1
2 to 4	0.3	±0.1
4 to 6	0.5	±0.1
6 to 8	0.7	±0.1
8 to 10	0.9	±0.1
9 to 10	0.95	±0.05

and, perhaps, snow depth effects. The first and second principal components are not included as they are effectively already used in the conventional SSM/I algorithms, although the principal components are coordinate transforms of the information used in the conventional SSM/I algorithms. The spectral gradient ratio in the NASA Team algorithm, for example, is directly comparable to the second principal component. The SI algorithm will ensure that if these low variance principal components do not add to the overall ability of the algorithm to predict ice concentration in a region, then they will not be used (the coefficients will be very small). Note that these components contain information from the 85-GHz channel, and so will be sensitive to atmospheric influences as well as thin ice cover and, perhaps, snow cover.

Although these so-called "partial" ice charts are not routinely retained at the NIC, for the purposes of this study, partial ice charts were collected for the Barents Sea for the period of November 2, 1998 to December 14, 1998. In order to make use of the NIC partial ice charts, the ice chart data had to be converted, in most cases, from a range of concentrations to a single concentration value. Ice analysts normally enter a range of ice concentrations for each region of the ice chart as they are not always able to determine ice concentrations to the nearest tenth. The conversions were carried out as shown in Table III.

In addition, the data were converted to a raster format, and polar stereographic projection consistent with the SSM/I antenna temperature data was provided by the NOAA NCEP (R. Grumbine, personal communication). In the partial ice charts, ice concentrations were indicated by a number between 0 and 100 (indicating ice concentration as a percentage of complete ice cover), and areas of no data were indicated by a value of −99. Implicit in this mapping procedure is the need to reduce the resolution of the ice chart data. This is achieved by a local averaging of the concentrations provided in the ice charts.

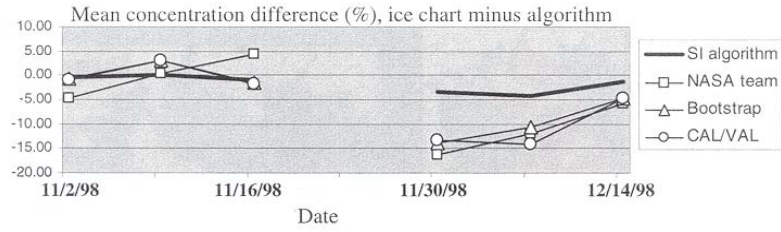

Fig. 10. Mean difference between ice concentrations from algorithms and partial ice charts as a function of date for the Barents Sea. A positive value indicates that the algorithm predicts more ice than the ice chart.

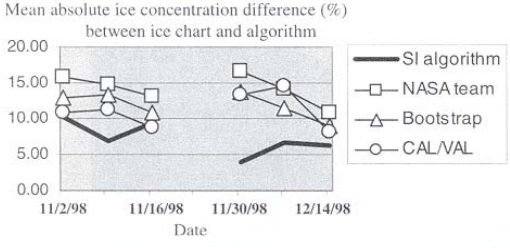

Fig. 11. Mean absolute difference between ice concentrations from algorithms and partial ice charts as a function of date for the Barents Sea.

Fig. 9 shows the results of applying the SI and conventional algorithms to the Barents Sea, together with the partial ice chart used to tune the SI algorithm. The subtle differences between the algorithms are difficult to see, but the figure does show that the SI algorithm is stable outside the area for which data were available to generate the coefficients, and this was found to be the case for all the examples that were investigated. Statistics have been derived to describe the differences between the partial ice charts and the SI and conventional algorithms. These statistics are illustrated in Figs. 10 and 11. Fig. 10 shows the mean difference between the partial ice chart and each algorithm as a function of date. This shows that the SI algorithm is successful in reducing much of the sometimes significant difference between the conventional algorithms and the partial ice charts derived from RADARSAT data and cloud-free AVHRR and OLS data. The mean difference reduces from a range of -15 to $+5\%$ for the conventional algorithms to a range of between -5 and 0% for the SI algorithm. In general, the mean difference between the SI algorithm and the partial ice chart is 30% (or less) of that of the conventional algorithms. Note that there is perhaps more similarity between the three conventional SSM/I algorithms than differences. In early November, the algorithms all predict mean ice concentrations within 5% of the partial ice chart (and thus within the error range of the ice chart). In late November and December, the three conventional algorithms show a significantly smaller mean concentration than that of the partial ice chart.

Fig. 11 shows the mean absolute difference between the ice chart and each algorithm as a function of date. It can be seen in this case that the SI algorithm reduces the mean absolute difference from around the typical 10–15% of the conventional algorithms to between 5–10%. These differences are quite large and reflect the fact that the test region includes the marginal ice zone. It must also be remembered, however, that there is imperfect matching in time between the SSM/I data and the partial ice chart source data (generally a difference of up to 36 h). If the ice charts were made available on a daily basis, as is the aim at the NIC for some regions [3], then this would improve the match between the SSM/I data and the ice charts, and at the same time, would improve the performance of the SI algorithm.

It is useful to consider the stability of the coefficients calculated in the SI algorithm from week to week. The variance due to the partial concentrations of the principal components (as opposed to the NASA Team and Bootstrap partial concentrations) varies from a minimum of 7.64% to a maximum of 48.5% and is not particularly consistent from week to week. This level of inconsistency from week to week probably reflects the sensitivity of these components to atmospheric influences as well as thin ice cover. The SI algorithm will effectively make use of the principal components as a form of weather filter as well as a method of detecting thin ice. The sensitivity to atmospheric conditions that is undoubtedly a feature of these principal components, largely as a result of the 85-GHz channel, means that the SI algorithm will modify the coefficients from week to week depending on changes in synoptic weather conditions.

Overall, these results indicate that the SI algorithm is successful at tuning an SSM/I algorithm to ice chart data. Furthermore, the technique does not exhibit any obvious "unstable" behavior within the "no data" region of the partial ice chart (in other words, outside the area that was used to calculate the model coefficients). However, there are certain limitations to the technique and its evaluation reported here that need to be made explicit.

1) There has been no independent verification of the performance of the technique. It has been shown that the technique is successful at closing the gap between ice chart-derived ice information and SSM/I-derived ice information. The SI algorithm will only be as good as the quality of the ice chart.
2) The SI algorithm will become less stable as the proportion of the area in the partial ice chart labeled "no data" increases.
3) The performance of the SI algorithm depends on a good match in time between the concentrations derived from the cloud-free visible and SAR data. If the data that are used to generate concentrations in the partial ice chart are spread over a period of more than perhaps two consecutive days, then the performance of the algorithm will

degrade significantly. In the ideal situation, the algorithm would be used in helping to compile daily ice charts where there is a reliably close match in time between the ice chart and the SSM/I data. With the current weekly ice charts from NIC, there will be some error introduced into the SI algorithm by using SSM/I data recorded on one day.

4) The technique has not been tested on summer data, although it is expected that it will compensate to some degree for the sensing difficulties associated with summer data. The tie point associated with summer sea ice will be closer to that of open water. The algorithm will use the ice chart to associate these modified tie points to 100% ice cover, and will generate the linear mapping coefficients accordingly.
5) The SSM/I data can provide a precision in ice concentrations that the NIC ice charts lack. This is because they provide precise calculations of ice concentration (even if they are subject to, in some cases, gross errors, as discussed earlier).

V. Conclusions and Recommendations

The SI algorithm is a practical technique for use by the operational ice community and generates significantly improved performance over conventional algorithms when referenced to ice chart information. In addition, the technique is practical for near-real time ice monitoring. The algorithm represents a new approach to the use of SSM/I data for operational ice monitoring, and has the following advantages:

1) automated tuning of the algorithm to the date and region of interest;
2) no manual intervention required (beyond production of the ice chart itself, which is an existing procedure);
3) some correction for sensing biases such as underestimation of thin ice and weather effects;
4) flexibility in terms of the attributes used in the program (new attributes can be added to increase the model fit achieved by the algorithm).

The disadvantage of the algorithm is that it is only as good as the ice chart from which it is derived. Any biases in the analysis procedure will be inherited by the algorithm. This algorithm will tend to reinforce any such biases by trying to include them through the linear coefficients. Thus, it is important that the analysis procedure is carefully monitored to minimize such effects.

The following recommendations are made.

1) The algorithm should be tested with different attributes to see whether there are systematic differences in the level of performance.
2) The algorithm should be tested for summer conditions when it is expected that algorithm will have some success in compensating for the more severe underestimation of ice concentrations by conventional SSM/I algorithms.

Operational implementation and testing of the algorithm is planned at the NIC.

Acknowledgment

The author wishes to thank M. R. Keller, A. Trakic, K. Dedrick, C. Bertoia, and P. Seymour of the U.S. National Ice Center, Washington, DC. SSM/I data for the study were kindly provided by R. Grumbine and P. Chang of the National Oceanographic and Atmospheric Administration, Washington, DC. He also wishes to thank J. Haferman, Fleet Numerical Modeling and Oceanography Center, Monterey, CA, for providing ice concentration products generated using the CAL/VAL and NASA Team algorithms. This research was carried out as part of the UCAR Visiting Scientist Program.

References

[1] A. Nicol and I. Allison, "The frozen skin of the Southern Ocean," *Amer. Scientist*, vol. 85, pp. 426–439, 1986.
[2] F. D. Carsey, Ed., "Microwave remote sensing of sea ice," in *Geophysical Monograph*. Washington, DC: Amer. Geophys. Union, vol. 68, ch. 1, pp. 1–6.
[3] C. Bertoia, J. Falkingham, and F. Fetterer, "Polar SAR data for operational sea ice mapping," in *Analysis of SAR Data of the Polar Oceans*, C. Tsatsoulis and R. Kwok, Eds, Berlin, Germany: Springer-Verlag, 1998.
[4] K. Steffen, J. Key, D. Cavalieri, J. Comiso, P. Gloersen, K. St. Germain, and I. Rubinstein, "The estimation of geophysical parameters using passive microwave algorithms," in *Microwave Remote Sensing of Sea Ice, Geophysical Monograph*, F. D. Carsey, Ed. Washington, DC: American Geophysical Union, 1992, vol. 68, ch. 1, pp. 1–6.
[5] D. Benner, J. Walsh, K. Dedrick, W. Chapman, and K. Partington, "Analysis of Digital Climatology Data for the Arctic Ocean," , submitted for publication.
[6] A. Cheng and R. Preller, "The Development of an Ice-Ocean Coupled Model in the Northern Hemisphere," Naval Research Laboratory, Washington, DC, Tech. Rep. NRL/FR/7322-95-9627, 1996.
[7] D. Cavalieri, P. Gloersen, and W. J. Campbell, "Determination of sea-ice parameters with the NIMBUS 7 SMMR," *J. Geophys. Res.*, vol. 89, no. D4, pp. 5355–5369, 1984.
[8] D. Cavalieri, "A microwave technique for mapping thin sea-ice," *Jnl. Geophys. Res.*, vol. 99, no. C6, pp. 12 561–12 572, 1994.
[9] J. Comiso, "Characteristics of arctic winter sea-ice from satellite passive microwave and infrared observations," *J. Geophys. Res.*, vol. 91, no. C1, pp. 975–994, 1986.
[10] J. R. Hollinger, R. Lo, G. Poe, R. Savage, and J. Pierce, "Special Sensor Microwave/Imager Calibration/Validation," Naval Research Laboratory, Washington, DC, Final Rep., 1991.
[11] D. Smith, "Extraction of winter total sea-ice concentration in the Greenland and Barents seas from SSM/I," *Int. J. Remote Sensing*, vol. 17, pp. 2625–2646, 1996.
[12] D. Cavalieri, J. P. Crawford, M. R. Drinkwater, D. T. Eppler, L. D. Farmer, R. R. Jentz, and C. C. Wackerman, "Aircraft active and passive microwave validation of sea-ice concentration from the defense meteorological satellite program, special sensor microwave/imager," *J. Geophys. Res.*, vol. 96, no. C12, pp. 21 989–22 008, 1991.
[13] J. Comiso, D. Cavalieri, C. Parkinson, and P. Gloersen, "Passive microwave algorithms for sea ice concentration: A comparison of two techniques," *Remote Sens. Environ.*, vol. 60, pp. 357–384, 1997.
[14] K. Emery, C. Fowler, and J. Maslanik, "Arctic sea ice concentrations from special sensor microwave imager and advanced very high resolution radiometer satellite data," *J. Geophys. Res.*, vol. 99, no. C9, pp. 18 329–18 342, 1994.
[15] K. Steffen and A. Schweiger, "NASA team algorithm for sea ice concentration retrieval from defense meteorological satellite program microwave imager: Comparison with landsat satellite imagery," *J. Geophys. Res.*, vol. 96, no. C12, pp. 21 971–21 987, 1991.
[16] A. W. Lohanik, "Some observations of established snow cover on saline ice and their relevance to microwave remote sensing," in *Proc. W. F. Weeks Symp.: Sea Ice Properties and Processes, CRREL Monograph*, S. F. Ackley and W. F. Weeks, Eds., Hanover, NH, 1990, pp. 61–67.
[17] T. C. Grenfell, "Surface microwave passive microwave observations of sea-ice in the Bering and Greenland seas," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-24, pp. 378–382, May 1986.

[18] D. Thomas, "Arctic sea-ice signatures for passive microwave algorithms," *J. Geophys. Res.*, vol. 98, no. C6, pp. 10 037–10 052, 1993.
[19] H. Kitahara, K. Okada, and S. Nishimura, "Improvement of SSM/I Sea Ice Concentration Algorithm for the Sea of Okhotsk," , submitted for publication.
[20] C. Oelke, "Atmospheric signatures in sea-ice concentration estimates from passive microwave sensors: Modeled and observed," *Int. J. Remote Sensing*, vol. 18, pp. 1113–1136, 1997.
[21] E. Svendsen, C. Matzler, and T. C. Grenfell, "A model for retrieving total sea-ice concentration from a spaceborne dual-polarized passive microwave instrument operating near 90 GHz," *Int. J. Remote Sensing*, vol. 8, pp. 1479–1487, 1987.
[22] D. Lubin, C. Garrity, R. Ramseier, and R. Whritner, "Total sea-ice concentration retrieval from the SSM/I 85.5 GHz channels during the Arctic summer," *Remote Sens. Environ.*, vol. 62, pp. 63–75, 1997.
[23] T. Markus and D. Cavalieri, "A Revised NASA Team Sea Ice Algorithm," , submitted for publication.
[24] J. Key, J. Maslanik, and A. J. Schweiger, "Classification of merged AVHRR and SMMR Arctic data with neural networks," *Photogramm. Eng. Remote Sensing*, vol. 55, pp. 1331–1338, 1989.
[25] S. G. Beaven and S. P. Gogineni, "Fusion of satellite SAR with passive microwave data for sea ice remote sensing," in *Analysis of SAR Data of the Polar Oceans*, C. Tsatsoulis and R. Kwok, Eds, Berlin, Germany: Springer-Verlag, 1998.
[26] D. Thomas and A. Rothrock, "Blending sequential scanning multi-channel microwave radiometer and buoy data into a sea-ice model," *J. Geophys. Res.*, vol. 94, no. C8, pp. 10 907–10 920, 1989.
[27] M. Wensnehan, G. A. Maykut, T. C. Grenfell, and D. P. Winebrenner, "Passive microwave remote sensing of thin sea ice using principal component analysis," *J. Geophys. Res.*, vol. 98, no. C7, pp. 12 453–12 468, 1993.
[28] T. C. Grenfell, D. G. Barber, A. K. Fung, A. J. Gow, K. C. Jezek, E. J. Knapp, S. V. Nghiem, R. G. Onstott, D. K. Perovich, C. S. Roesler, C. T. Swift, and F. Tanis, "Evolution of electro-magnetic signatures of sea ice from initial formation to establishment of thick first-year ice," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1642–1654, Sept. 1998.
[29] D. A. Rothrock, D. R. Thomas, and A. S. Thorndike, "Principal component analysis of satellite passive microwave data over sea ice," *J. Geophys. Res.*, vol. 93, no. C3, pp. 2321–2332, 1988.

Kim C. Partington received the B.A. degree in geography from the University of Cambridge, U.K., in 1983, and the Ph.D. degree in geophysics from University College, London, U.K., in 1988.

He was a Research Scientist with GEC-Marconi Research Centre, Chelmsford, U.K., for eight years and was then with the Alaska SAR Facility, University of Alaska, Fairbanks, and the U.S. National Ice Center, Suitland, MD, where he was Senior Scientist. He is currently Manager of the Polar Research Program, NASA Headquarters, Washington, DC.